

Aggregate structure of hydroxyproline-rich glycoprotein (HRGP) and HRGP assisted dispersion of carbon nanotubes

Ben Wegenhart · Li Tan · Michael Held ·
Marcia Kieliszewski · Liwei Chen

Published online: 1 August 2006
© to the authors 2006

Abstract Hydroxyproline-rich glycoproteins (HRGP) comprise a super-family of extracellular structural glycoproteins whose precise roles in plant cell wall assembly and functioning remain to be elucidated. However, their extended structure and repetitive block co-polymer character of HRGPs may mediate their self-assembly as wall scaffolds by like-with-like alignment of their hydrophobic peptide and hydrophilic glycopeptide modules. Intermolecular crosslinking further stabilizes the scaffold. Thus the design of HRGP-based scaffolds may have practical applications in bionanotechnology and medicine. As a first step, we have used single-molecule or single-aggregate atomic force microscopy (AFM) to visualize the structure of YK20, an amphiphilic HRGP comprised entirely of 20 tandem repeats of: Ser-Hyp₄-Ser-Hyp-Ser-Hyp₄-Tyr-Tyr-Tyr-Lys. YK20 formed tightly aggregated coils at low ionic strength, but networks of entangled chains with a porosity of ~0.5–3 μm at higher ionic strength. As a second step we have begun to design HRGP-carbon nanotube composites. Single-walled carbon nanotubes (SWNTs) can be considered as seamless cylinders rolled up from graphene sheets. These unique all-carbon structures have extraordinary aromatic and hydrophobic properties and form aggregated bundles due to strong inter-tube van der Waals interactions. Sonicating aggregated SWNT bundles with aqueous YK20 solubilized them presumably by interaction with the repetitive, hydrophobic, Tyr-rich peptide modules

of YK20 with retention of the extended polyproline-II character. This may allow YK20 to form extended structures that could potentially be used as scaffolds for site-directed assembly of nanomaterials.

Keywords Hydroxyproline-rich glycoprotein · Carbon nanotube · Nano assembly

Introduction

Hydroxyproline-rich glycoproteins (HRGPs) comprise a superfamily of extra-cellular structural proteins expressed in plant cell walls and extracellular matrix during normal development and in response to stress [1, 2]. HRGPs are extended macromolecules consisting of small repetitive peptide and glycopeptide motifs. While the peptide motifs often contain hydrophobic tyrosine residues, the glycopeptide motifs result from a combination of post-translational modifications unique to plants, namely proline hydroxylation and subsequent hydroxyproline (Hyp) glycosylation. The precise oligosaccharides or polysaccharide decoration pattern is driven by a sequence-dependent glycosylation code [2–4]. The key to this glycosylation code is Hyp contiguity: contiguous Hyp residues direct the addition of small arabinooligosaccharides to Hyp, while clustered non-contiguous Hyp residues direct the addition of larger complex hetero-polysaccharides. The addition of short oligosaccharides to Hyp residues locks the contiguous Hyp-rich glycopeptide motifs into an extended, left-handed polyproline-II helix conformation and thus results in rigid hydrophilic regions. In contrast, regions that lack contiguous Hyp remain flexible while subsequent addition of long

B. Wegenhart · L. Tan · M. Held · M. Kieliszewski ·
L. Chen (✉)
Department of Chemistry and Biochemistry, Ohio
University, Athens, Ohio 45701, USA
e-mail: chenl1@ohio.edu

polysaccharide to clustered non-contiguous Hyp residues promotes an extended random coil conformation [3].

Some HRGPs also contain hydrophobic, tyrosine-rich peptide motifs that function in intra- and intermolecular crosslinking. Indeed, using a synthetic gene approach we recently expressed in tobacco cells a simple arabinosylated HRGP analog containing 20 tandem repeats of the sequence: Ser-Hyp₄-Ser-Hyp-Ser-Hyp₄-Tyr-Tyr-Tyr-Lys, designated YK20, and demonstrated that YK20 was extensively crosslinked enzymically *in vitro* to give tyrosine-based intermolecular crosslinks [5]. This indicated that YK20 rapidly aligns itself for subsequent intermolecular crosslinking and raised questions about the aggregate structure of YK20 that drives this self-assembly, the networks that arise and whether or not their properties can be tailored for specific applications.

Here we report the first visualization of an YK20 ‘network’ by the single-molecule or single-aggregate imaging approach using atomic force microscopy (AFM), the first such characterization for any HRGP. We also noted that YK20, an amphiphilic molecule, interacted with single-walled carbon nanotubes (SWNTs) and dispersed SWNTs in aqueous solutions, which raised the possibility that SWNT-YK20 complexes might be exploited to yield templates for the assembly of high order structures.

Experimental methods

YK20 synthetic gene construction, plant cell transformation and YK20 glycoprotein isolation

A synthetic gene, *YK20-EGFP*, encoding 20 tandem repeats of the protein sequence Ser-Pro₄-Ser-Pro-Ser-Pro₄-Tyr-Tyr-Tyr-Lys fused to the gene for the enhanced green fluorescent protein (*EGFP*; Clontech) was constructed, tobacco cells (Bright Yellow 2) transformed, and the YK20 glycoprotein isolated after EGFP removal, all as previously described [5].

Dispersion of SWNTs in YK20 solutions

About 2 mg of HiPCO carbon nanotubes (carbon nanotechnology, Inc.) were added to a solution of 1 mg of YK20 in 1 mL of water. The mixture was vigorously sonicated using a sonication probe for an hour with ~5W power. The resulting suspension was then centrifuged at 14,000 g for an hour. The supernatant contained a solution of SWNT-YK20 complexes.

Atomic force microscopy

1 mg/mL solutions of YK20 were mixed in a 1:1 ratio with solutions of MgCl_2 , and then 20 μL of the mixture was spin-coated onto freshly cleaved mica for 50 s at 4000 rpm. Samples with high salt concentration had to be rinsed briefly with water and dried with nitrogen gas before they could be imaged. These samples were analyzed with an Alpha-SNOM atomic force microscope (Witech instrument Inc. Ulm, Germany) in the acoustic mode. SWNT-YK20 complexes were spin-coated onto freshly cleaved mica for 50 s at 2000 rpm. These samples were analyzed with an MFP-3D microscope (Asylum Research, Santa Barbara, CA) in AC mode. Si probes with spring constants of $\sim 4 \text{ N/m}$ and resonance frequencies of $\sim 75 \text{ KHz}$ (NSC18/AIBS, Micromasch, Estonia) were used for AFM imaging.

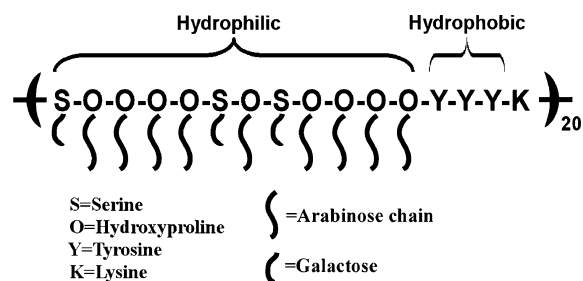
Absorption and circular dichorism (CD) spectroscopy

UV-visible absorption spectra were obtained on Agilent 8453 UV-vis spectrophotometer (Agilent Technologies, Palo Alto, CA) and the CD spectra were recorded on a Jasco-715 spectropolarimeter (Jasco Inc., Easton, MD). Spectra were averaged over two scans with a bandwidth of 1 nm, and step resolution was 0.1 nm. All spectra were reported in terms of mean residue ellipticity within the 180–250 nm region using a 1 mm path length. Samples of YK20 and SWNT-YK20 complexes were dissolved in water at a final protein concentration of 100 $\mu\text{g/mL}$.

Results and discussion

Aggregate structure of YK20

The YK20 primary amino acid sequence is shown in Scheme 1 along with glycan assignments. The genetically engineered HRGP contains 20 tandem repeats each containing a long hydrophilic stretch of monogalactosylated serine and arabinosylated hydroxyproline



Scheme 1 Amino acid sequence of YK20

residues followed by a short hydrophobic block of three tyrosine residues and a positively charged lysine residue. YK20 proteins were deposited from a solution to freshly cleaved mica surfaces for AFM imaging. When dissolved in a solution of low ionic strength, YK20 yielded large aggregates a few micrometers in diameter, however the higher ionic strength solution produced open networks of entangled fibrils (Fig. 1).

The single molecule or single aggregate imaging approach using AFM provides direct visualization of biological macromolecules [6–8]. It is a new method in structural biology that complements traditional crystallography and nuclear magnetic resonance methods [9–12] and is particularly well-suited for HRGPs, which are extended rods, highly glycosylated and possess too much heterogeneity in high order structures for X-ray crystallography or NMR techniques.

Since the single-molecule approach is a surface bound imaging technique, it is important to make sure the snapshots imaged on surfaces represent the equilibrium structures in solutions and yield information that agree with conventional biochemical studies. There exists a large parameter window of sample deposition conditions on mica for long linear molecules, such as double-stranded DNA, in which the molecular configuration on 2D surfaces accurately reflects the configuration of free molecules in 3D solutions [9, 10]. Therefore, we chose mica as the substrate for AFM imaging of YK20 in order to retain its native structure on substrates.

Four interaction forces likely contribute to YK20 homophilic interactions and the formation of aggregates. Firstly, hydrophobic interactions between the repetitive tyrosine blocks; secondly, interactions between positively charged lysine residues and the negatively charged C-terminus; thirdly, lysine residues may also interact with the aromatic rings of the tyrosine residues through cation- π interactions [13, 14]; and finally, the Mg^{2+} ions undoubtedly promote homophilic associations between the extensively glycosylated Ser-Hyp₄ glycomodules as already demonstrated for Ca^{2+} ions (Tan, Sulaiman, Tees and Kieliszewski, unpublished data). At high ionic strength, the electrostatic, cation- π and hydrophobic interactions are screened by the redistribution of ions in solution and thus the condensed aggregates opened up and displayed the random networks of linear fibrils. Since the polyproline-II helix is a left-handed helix with about 3 residues per turn and a pitch of 9.4 angstrom, the length of YK20 is only ~100 nm (320 amino acids) and the entangled network clearly consists of multiple molecules. The height of the linear fibril ranges from less than 1 nm to about 4 nm, thus the open aggregates are likely individual helices or at most only a few associated YK20 molecules.

The observed aggregation agrees with earlier work demonstrating the very rapid in vitro crosslinking of YK20 by a plant peroxidase [5], which indicated YK20 monomers align their tyrosine residues for subsequent intermolecular crosslinking. The ability of YK20 to

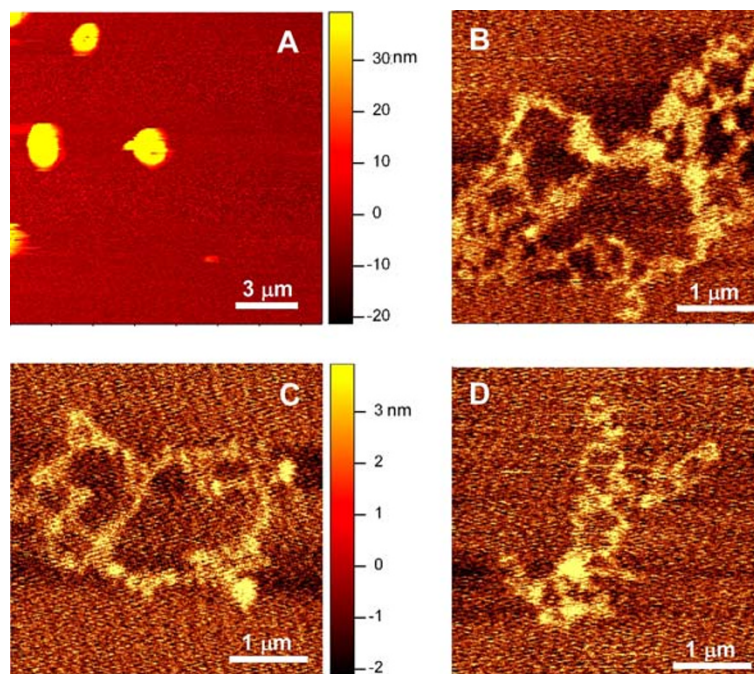


Fig. 1 Atomic force microscopy images of YK20 aggregate structures. (A): YK20 on mica deposited from a solution with 12 mM $MgCl_2$; (B), (C), and (D): YK20 on mica deposited from 70 mM $MgCl_2$ solutions

align the hydrophobic tyrosine blocks and form aggregated structures raises the possibility that YK20 might interact with hydrophobic non-biological materials such as carbon nanotubes.

YK20 assisted dispersion of SWNTs

One reason for studying the interactions between YK20 and hydrophobic materials comes from our search of surfactant for SWNTs. SWNTs are a family of nanomaterials whose structure can be regarded as seamless hollow cylinders rolled up from graphene sheets [15, 16]. SWNTs have not only inspired much interest in fundamental sciences due to their unique all-carbon one-dimensional structure, but also showed great potential in a wide variety of applications ranging from composite materials, molecular electronics, and chemical and biological sensors, to electrochemical cells and fuel cells for alternative energy solutions [17–19]. As-produced SWNTs form closely packed bundles due to the strong inter-tube van der Waals interactions and hydrophobic interactions in aqueous environments. But most applications require well-dispersed SWNT systems in order to take advantage of the unique properties. Many surfactants such as lipids, sugars, proteins, DNA, commodity polymers, and designed polymers have been used to facilitate the dispersion of SWNTs [20–26]. Given the amphiphilicity of YK20, we examined the SWNT dispersing properties of YK20 in solution.

Figure 2A shows solutions of SWNT-YK20 complexes obtained after vigorous sonication and extended centrifugation to pellet non-complexed insoluble nanotubes. The microfuge tube at the far left shows SWNTs solubilized in a 1 mg/mL solution of YK20 while the microfuge tube to the right shows the 10-fold dilution of the SWNT-YK20 solution. The UV-vis absorption spectrum (Fig. 2B) shows peaks that agree with the first van Hove transitions from metallic tubes (~ 400 – 600 nm) and the second van Hove transitions from semiconducting tubes (~ 500 – 900 nm) in published literature [20, 24, 26]. The peaks are not as sharp as seen in other surfactant micelles, such as ssDNA or designed polysoap. This suggests that small bundles of SWNTs may still exist in the solution.

AFM images of the SWNT-YK20 complexes corroborate the observations above. While absolutely no individual or small bundles of SWNTs could be found in solution without YK20 treatment, Fig. 3 shows that in the presence of YK20, the majority of the SWNTs are individually dispersed with tube heights about 0.7–2 nm and lengths ~ 500 – 1500 nm. Small bundles of about 6–12 nm in diameter are also seen. Furthermore, there is no evidence of the large YK20 aggregates featured in Fig 1, presumably because they were disbanded through preferred interaction of YK20 monomers with SWNTs.

The mechanism by which YK20 facilitates the dissolution of SWNTs in aqueous solution is suggested by its structure. As shown in scheme 1, YK20 consists

Fig. 2 (A) Photographs of SWNT-YK20 complexes in solutions, and (B) UV-vis absorption spectrum

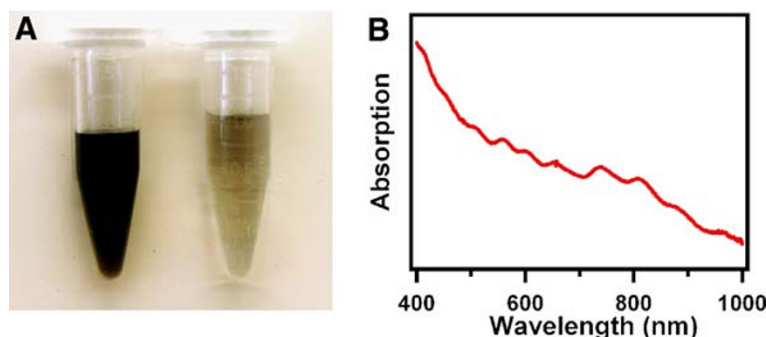
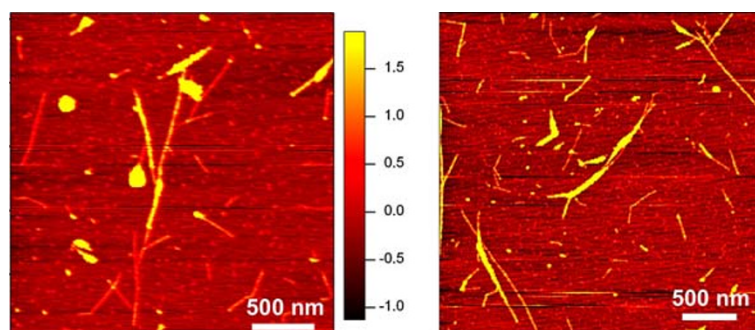


Fig. 3 Atomic force microscopy images of SWNT-YK20 complexes deposited on mica. The two images share the same color map



of alternating hydrophilic and hydrophobic blocks and effectively is an amphiphilic block-copolymer. Amphiphilic macromolecules such as designed peptides [22] and linear DNA molecules [20] disperse SWNTs by interacting extensively with the nanotube side walls through the hydrophobic effects. Similarly, YK20 molecules probably coat the SWNTs through interactions involving the Tyr-Tyr-Tyr hydrophobic segments and solvate the complex through the blocks of hydrophilic amino acids (Ser and Hyp) and the abundant glycans that bind water. The details regarding the YK20 configuration around SWNTs demands atomically resolved microscopy techniques and will be pursued in future studies. However, circular dichroism spectroscopy of YK20 alone and YK20 in complex with SWNTs (Fig. 4) suggested YK20 underwent significant conformational changes upon SWNT complexation.

SWNT induced changes in YK20 structures

While the strong interactions between YK20 and SWNTs help to disperse SWNTs in water, they may also simultaneously influence the structure of YK20. Shown in the Fig. 4 are the circular dichroism (CD) spectra of YK20 and SWNT-YK20 complexes. The pronounced features in the spectra, a minimum at around 205 nm and a maximum at around 223 nm, are associated with the left-handed polyproline II helix [3]. The green dashed line, whose intensity at both the minimum and the maximum are the same as pure YK20, is the spectrum of SWNT-YK20 complexes multiplied by a factor of 2.25. Thus, the binding of SWNTs imposes some other conformations of YK20, and the polyproline II content in the secondary structure of the protein is reduced to about 45%. Moreover, the CD spectrum of the complexes has a shift of about 2–3 nm to longer wavelengths, which indicates the

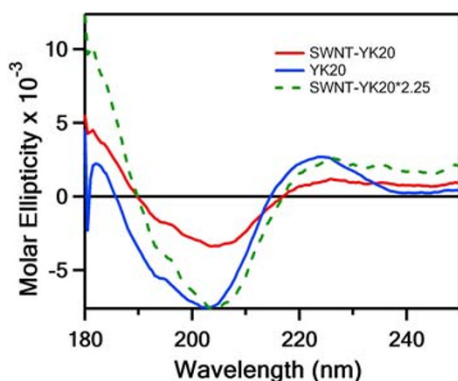


Fig. 4 Circular dichroism spectra of YK20 and SWNT-YK20 complexes. The protein concentration in both solutions is 100 $\mu\text{g/mL}$

YK20 secondary structure is also qualitatively different from that in free YK20 solution.

More interesting is the clear change in the aggregate structure shown in AFM images. In the absence of SWNTs, YK20 molecules interact among themselves and form either large condensates of hundreds of nanometers in diameter or open networks of more than 1 μm in size (Fig. 1). After complexing with SWNTs, neither of these aggregate structures is found (Fig. 2) and YK20 molecules most likely form extended structures along with SWNTs. The more ordered extended structure opens up new pathways to hierarchy assembly of nanomaterials. In combination with complete control over the primary sequence via genetic engineering, the extended HRGP structure in three-dimensional space may be used as scaffolds and templates for attaching other nano-building blocks at specific sites.

Conclusion

We have demonstrated that YK20, a genetically engineered HRGP, forms closely aggregated coils in low ionic strength solutions, and random networks of entangled chains at high ionic strength conditions. The hydrophobic segments of YK20 may interact with highly hydrophobic SWNTs and disperse them in aqueous solutions. The dispersion of SWNTs is an important step towards solution processing and applications of this unique nanomaterial. More interestingly, it helps to stabilize extended and ordered aggregate structures of YK20, which is not favored in pure protein solutions. The YK20 proteins are stretched along the side walls of SWNT and result in significantly different CD spectra of the protein. SWNT induced extended structure of HRGPs could potentially be used as scaffolds for site-directed assembly of nanomaterials.

Acknowledgments B. W. thanks the Ohio University PACE (Program to Aid Career Exploration) for financial support. This project was supported by grants from the Ohio University NanobioTechnology Initiative (NBTI), the Herman Frasch Foundation (526-HF02), and the United States Department of Agriculture (2004-34490-14579).

References

1. M.J. Kieliszewski, *Phytochemistry* **57**(3), 319–323 (2001)
2. M.J. Kieliszewski, E. Shpak, *Cell. Mol. Life Sci.* **58**(10), 1386–1398 (2001)
3. E. Shpak, E. Barbar, J.F. Leykam, M.J. Kieliszewski, *J Biol Chem* **276**(14), 11272–11278 (2001)

4. E. Shpak, J.F. Leykam, M.J. Kieliszewski, *Proc. Nat. Acad. Sci. USA* **96**(26), 14736–14741 (1999)
5. M.A. Held, L. Tan, A. Kamyab, M. Hare, E. Shpak, M.J. Kieliszewski, *J. Biol. Chem.* **279**(53), 55474–55482 (2004)
6. C. Bustamante, C. Rivetti, D.J. Keller *Curr. Opin. Struct. Biol.* **7**(5), 709–716 (1997)
7. H.G. Hansma, K. Kasuya, E. Oroudjev, *Curr. Opin. Struct. Biol.* **14**(3), 380–385 (2004)
8. H.G. Hansma, L.I. Pietrasanta, I.D. Auerbach, C. Sorenson, R. Golan, P.A. Holden, *J. Biomater. Sci. Polym. Ed.* **11**(7), 675–683 (2000)
9. L.W. Chen, K.A. Haushalter, C.M. Lieber, G.L. Verdine, *Chem. Biol.* **9**(3), 345–350
10. C. Bustamante, C. Rivetti *Annu. Rev. Biophys. Biomol. Struct.* **25**, 395–429 (1996)
11. J.H. Hafner, C. Cheung, A.T. Woolley, C.M. Lieber *Prog. Biophys. Mol. Biol.* **77**(1), 73–110 (2001)
12. A.T. Woolley, C. Guillemette, C.L. Cheung, D.E. Housman, C.M. Lieber *Nat. Biotechnol.* **18**(7), 760–763 (2000)
13. D.A. Dougherty *Science* **271**(5246), 163–168 (1996)
14. J.D. Schmitt, C.G.V. Sharples, W.S. Caldwell, *J. Med. Chem.* **42**(16), 3066–3074 (1999)
15. T.W. Odom, J.L. Huang, P. Kim, C.M. Lieber, *J. Phys. Chem. B* **104**(13), 2794–2809 (2000)
16. R. Saito, G. Dresselhaus, M.S. Dresselhaus, *Physical Properties of Carbon Nanotubes* (Imperial College Press, London, 1998)
17. P. Avouris, *Acc. Chem. Res.* **35**(12), 1026–1034 (2002)
18. R.H. Baughman, A.A. Zakhidov, de W.A. Heer *Science* **297**(5582), 787–792 (2002)
19. M. Ouyang, J.L. Huang, C.M. Lieber, *Acc. Chem. Res.* **35**(12), 1018–1025 (2002)
20. M. Zheng, A. Jagota, E.D. Semke, B.A. Diner, R.S. McLean, S.R. Lustig, R.E. Richardson, N.G. Tassi, *Nat. Mater.* **2**(5), 338–342 (2003)
21. V.A. Sinani, M.K. Gheith, A.A. Yaroslavov, A.A. Rakhnyanskaya, K. Sun, A.A. Mamedov, J.P. Wicksted, N.A. Kotov, *J. Am. Chem. Soc.* **127**(10), 3463–3472 (2005)
22. G.R. Dieckmann, A.B. Dalton, P.A. Johnson, J. Razal, J. Chen, G.M. Giordano, E. Munoz, I.H. Musselman, R.H. Baughman, R.K. Draper, *J. Am. Chem. Soc.* **125**(7), 1770–1777 (2003)
23. D. Wang, W.X. Ji, Z.C. Li, L.W. Chen, *J. Am. Chem. Soc.* ASAP article (2006)
24. M.J. O’Connell, S.M. Bachilo, C.B. Huffman, V.C. Moore, M.S. Strano, E.H. Haroz, K.L. Rialon, P.J. Boul, W.H. Noon, C. Kittrell, J.P. Ma, R.H. Hauge, R.B. Weisman, R.E. Smalley, *Science* **297**(5581), 593–596 (2002)
25. C. Richard, F. Balavoine, P. Schultz, T.W. Ebbesen, C. Mioskowski *Science* **300**(5620), 775–778 (2003)
26. M. Zheng, A. Jagota, M.S. Strano, A.P. Santos, P. Barone, S.G. Chou, B.A. Diner, M.S. Dresselhaus, R.S. McLean, G.B. Onoa, G.G. Samsonidze, E.D. Semke, M. Usrey, D.J. Walls *Science* **302**(5650), 1545–1548 (2003)